

WHAT IS CLAIMED IS:

1                   1.       An apparatus for converting a bitstream representing frames of data  
2 encoded according to a first voice compression standard to a bitstream representing frames of  
3 data according to a second voice compression standard and/or within a single standard but to  
4 a different mode, comprising:

5                   a bitstream unpacking module for extracting one or more CELP parameters  
6 from a first voice codec bitstream frame;

7                   a parameter interpolator module coupled to the bitstream unpacking module,  
8 the parameter interpolator module interpolates between different frame sizes, subframe sizes,  
9 and/or sampling rates of a first voice codec and a second voice codec;

10                  a frame classification and rate determination module coupled to the parameter  
11 interpolator module, the frame classification and rate determination module produces a frame  
12 class and rate decision based one or more inputs from CELP parameters of the first voice  
13 codec, a type of bitstream frames of the first voice codec, and one or more external control  
14 commands of the second voice codec;

15                  a first-stage switching module coupled to the parameter interpolator module  
16 and the frame classification and rate determination module, the first-stage switching module  
17 being functioned to link the parameter interpolator module to a mapping module and being  
18 controlled by one or more outputs of the frame classification and rate determination module;

19                  a mapping module coupled to the parameter interpolator module through a  
20 switching module, the mapping module being adapted to map the one or more CELP  
21 parameters from the first voice codec to one or more CELP parameters of the second voice  
22 codec;

23                  a second-stage switching module coupled to the mapping module, the second-  
24 stage switching module being functioned to link the mapping module and a destination  
25 bitstream packing module and being controlled by one or more outputs of the frame  
26 classification and rate determination module;

27                  a destination bitstream packing module coupled to a mapping module through  
28 the second-stage switching module, the destination bitstream packing module being adapted  
29 to construct at least one destination output CELP frame based upon at least one CELP  
30 parameter from the second voice codec; and

31                  a controller coupled to the destination bitstream packing module, the  
32 interpolator module, the mapping module, the frame classification and rate determination

33 module and the bitstream unpacking module, whereupon the controller is adapted to oversee  
34 the operation of one or more of the modules, to receive instructions from one or more  
35 external applications, and to provide status information to one or more external applications.

1                   2.       The apparatus of claim 1 wherein the controller is a single controller or  
2 multiple controllers.

1                   3.       The apparatus of claim 1 wherein the interpolator module, the mapping  
2 module, the first stage switching module, the second switching module, and the destination  
3 bitstream packing module are combined into a single module.

1                   4.       The apparatus of claim 1 wherein the mapping module is a single  
2 module or multiple modules.

1                   5.       The apparatus of claim 1 wherein the frame classification and rate  
2 determination module is a single module or multiple modules.

1                   6.       The apparatus of claim 1, wherein the bitstream unpacking module  
2 comprises:  
3                   a bitstream processor, the bitstream processor extracts information in a first  
4 format from one or more compressed voice parameters in a first voice codec input frame;  
5                   an LSP decoding module coupled to the bitstream processor, the LSP  
6 decoding module outputs one or more LSP coefficients using at least the information from  
7 the compressed input frame of the first voice codec;  
8                   a decoding module coupled to the bitstream processor, the decoding module  
9 decodes the pitch information to output a pitch lag parameter and a pitch gain parameter from  
10 the compressed input frame of the first voice codec;  
11                  a fixed codebook decoding module coupled to the bitstream processor, the  
12 fixed codebook decoding module decodes the fixed codebook gain and fixed codevector  
13 indices to output a gain-adjusted fixed codebook vector;  
14                  an adaptive codeword decoding module coupled to the bitstream processor,  
15 the adaptive codeword decoding module decodes the pitch lag and pitch gain information to  
16 output an gain-adjusted adaptive codebook contribution vector;  
17                  an energy decoding module coupled to the bitstream processor, the frame  
18 energy decoding module decodes the frame energy or subframe energies; and

19 a shape decoding module coupled to the bitstream processor, the shape  
20 decoding module decodes the excitation shape.

1 7. The apparatus of claim 1, wherein the interpolator module comprises  
2 an LSP process, the LSP process converts one or more LSP coefficients of a  
3 source codec into one or more LSP coefficients of a destination codec when the source codec  
4 and destination codec have a different subframe size and/or frame size;  
5 an adaptive codebook process, the adaptive codebook process converts a pitch  
6 lag and a pitch gain from the source codec into a preliminary estimate of the pitch lag and  
7 pitch gain of the destination codec when said source codec and destination codec have a  
8 different subframe size and/or frame size;  
9 a fixed codebook process, the fixed codebook process converts a fixed  
10 codebook gain or frame energy or subframe energy from the source codec into a preliminary  
11 estimate of the fixed codebook gain or frame energy or subframe energy of the destination  
12 codec when said source codec and destination codec have a different subframe size and/or  
13 frame size; and  
14 a CELP parameter buffer, the CELP parameter buffer stores one or more  
15 CELP parameters that are not interpolated until a next first voice compression codec and a  
16 next second voice compression codec have a different subframe size and/or frame size.

1 8. The apparatus of claim 1, wherein the frame classification and rate  
2 determination module comprises:  
3 A classifier input parameter selector, for selecting which inputs will be used in  
4 the classification task;  
5 M sub-classifier modules which take as input the selected classification input  
6 parameters and output the frame class and rate decision;  
7 buffers to store past input parameters and past output values; and  
8 a final decision module that takes inputs from external commands and the  
9 output of the M sub-classifier module, and generates the frame class and rate decision.

1 9. The apparatus of claim 1, wherein the mapping module comprises:  
2 a mapping and tuning module, the mapping and tuning module outputs the one  
3 or more destination voice compression parameters; and  
4 a decision module, the decision module selects a voice compression parameter  
5 mapping strategy based upon a plurality of strategies.

1           10.     The apparatus of claim 9 wherein the plurality of strategies comprises:  
2           CELP parameter direct space mapping;  
3           analysis in excitation space  
4           analysis in filtered excitation space; and  
5           a combination of analysis in excitation space and filtered excitation space.

1           11.     The apparatus of claim 9, wherein the mapping and tuning module  
2 comprises:  
3           an spectral parameter mapping module that converts the source spectral  
4 parameters to encoded destination spectral parameters;  
5           a excitation mapping unit that takes source CELP excitation parameters  
6 including pitch lag, gain, and excitation vectors from the interpolation module to get encoded  
7 destination CELP excitation parameters.

1           12.     The apparatus of claim 11, wherein the excitation mapping unit  
2 comprises:  
3           a module of CELP parameters direct space mapping that produces encoded  
4 destination CELP parameters using analytical formula without any analysis or iterations;  
5           an excitation generator, the excitation generator outputs an excitation vector  
6 using at least the fixed codebook vector and the adaptive codebook vector, or at least the  
7 frame or subframe energy and a random noise generator, or at least the frame or subframe  
8 energy, shaping filter(s) and a random noise generator;  
9           a module of excitation domain mapping that produces encoded destination  
10 CELP parameters by searching in the excitation space;  
11           a module of filtered excitation domain mapping that produces encoded  
12 destination CELP parameters by searching adaptive closed-loop in filtered excitation space  
13 and fixed-codebook in filtered excitation space;  
14           a module of combined excitation domain mapping and filtered excitation  
15 domain mapping that produces encoded destination CELP parameters by searching some  
16 parameters in excitation space and some parameters in filtered excitation space.

1           13.     The apparatus of claim 11 wherein the excitation mapping unit may or  
2 may not include the process of modifying the signal to match an interpolated pitch delay  
3 contour.

1                   14.     The apparatus of claim 1, wherein the destination codec bitstream  
2     packing module comprises a plurality of frame packing facilities, each of the facilities being  
3     capable of adapting to a preselected application from a plurality of applications for a selected  
4     destination CELP coder, the selected destination CELP coder being one of a plurality of  
5     CELP coders.

1                   15.     The apparatus of claim 1, wherein the controller comprises:  
2                   a control unit which receives external instructions and controls each of the  
3     signal processing modules;  
4                   a status unit which sends transcoding information such as frame counts, error  
5     logs and other information to external modules upon request.

1                   16.     The apparatus of claim 7, wherein the interpolation module can select  
2     either linear interpolation or non-linear interpolation.

1                   17.     The apparatus of claim 7, wherein the CELP parameter buffer  
2     comprises:  
3                   an excitation vector buffer, the excitation vector buffer stores the  
4     reconstructed excitation vector that is to be mapped in the next subframe or frame;  
5                   an LSP coefficient buffer that stores the LSP coefficients before or after  
6     interpolation that are to be mapped in the next subframe or frame;  
7                   a buffer for other CELP parameters that stores the pitch lag, pitch gain,  
8     codebook gain and index before or after interpolation that are to be mapped in the next  
9     subframe or frame.

1                   18.     A method for transcoding a compressed voice bitstream from source  
2     codec to variable-rate destination codec, comprising:  
3                   processing a source codec input bitstream to unpack at least one or more voice  
4     parameters from an input bitstream in the case of CELP-based codecs the voice parameters  
5     include at least LSPs, pitch lag, adaptive codebook gain, fixed codebook gain, and fixed  
6     codevectors;  
7                   classifying a frame type of a destination codec from the one or more input  
8     parameters of the source codec;  
9                   determining the rate of the destination codec output from one or more input  
10    parameters of the source codec and external control commands;

11                   interpolating one or more of a plurality of unpacked voice parameters from a  
12 source codec format to a destination codec format if a difference exists between the frame  
13 size, subframe size, or sampling rate of the destination codec and the frame size, subframe  
14 size, or sampling rate of the source codec;  
15                   mapping the source CELP parameters to destination CELP parameters using a  
16 selected mapping strategy;  
17                   encoding the one or more CELP parameters for the destination codec; and  
18                   processing a destination bitstream by packing the one or more voice  
19 parameters for the destination codec.

1                   19.     The method of claim 18, wherein the processing of the source codec  
2 input packet comprises:  
3                   converting an input bitstream frame into information associated with one or  
4 more CELP parameters ;  
5                   decoding the information into one or more CELP parameters;  
6                   outputting the unquantized parameters to the interpolator.

1                   20.     The method of claim 18, wherein classifying the frame type of the  
2 destination codec comprises the steps of:  
3                   selecting classifier input parameters from one or more source codec CELP  
4 parameters and the source bit rate through a selection module;  
5                   using one or more external commands;  
6                   using previously stored state information;  
7                   performing frame classification according to pre-defined coefficients or rules;  
8 and  
9                   updating and storing the states for use in classifying future frames.

1                   21.     The method of claim 18, wherein determining the rate of a destination  
2 codec comprises the steps of:  
3                   selecting classifier input parameters from one or more source codec CELP  
4 parameters and the source bit rate through a selection module;  
5                   using one or more external commands;  
6                   using previously stored state information;  
7                   performing rate determination according to pre-defined coefficients or rules;  
8 and

9 updating the stored states for use in determining the rate of future frames.

1 22. The method of claim 20 wherein the pre-defined coefficients for frame  
2 classification are predetermined during a setup training process or construction process.

1 23. The method of claim 21 wherein the pre-defined coefficients for rate  
2 determination are predetermined during a setup training process or construction process.

1 24. The method of claim 18, wherein interpolating parameters comprises  
2 the steps of:

3 interpolating one or more of the LSP coefficients from the source codec to one  
4 or more LSP coefficients for the destination codec; and

5 interpolating CELP parameters other than the LSP coefficients from the source  
6 codec to CELP parameters for the destination codec.

1 25. The method of claim 24, further comprising:

2 converting the one or more LSP coefficients using a linear transform process.

1 26. The method of claim 18, wherein mapping the parameters further  
2 comprises the steps of:

3 mapping the source interpolated LSP coefficients to destination LSP  
4 coefficients and quantizing the destination LSP coefficients;

5 mapping the source interpolated excitation parameters to destination excitation  
6 parameters and quantizing the destination excitation parameters; and

7 selecting one of a plurality of CELP mapping strategies according to the  
8 control signal from the decision module.

1 27. The method of claim 26, wherein mapping the source interpolated  
2 excitation parameters further comprises the steps of:

3 reconstructing the source excitation signal if required by the mapping strategy;

4 filtering the reconstructed source excitation signal with a filter that accounts

5 for the differences between the quantized destination LP coefficients and quantised source LP  
6 coefficients to form the calibrated excitation vector;

7 transferring the calibrated excitation vector to another process; and

8 transferring the source excitation vector to the encoding process if the  
9 excitation vector does not require a calibration.

1                   28.     The method of claim 26 wherein the plurality of CELP mapping  
2 strategies includes:

3                   direct space mapping of the CELP parameters;  
4                   analysis in the excitation space;  
5                   analysis in the filtered excitation space; and  
6                   analysis in the combined excitation space and filtered excitation space.

1                   29.     A method of claim 27 where the reconstructed excitation signal may or  
2 may not be modified to match an interpolated delay contour.

1                   30.     The method of claim 28, wherein the step of direct space mapping of  
2 CELP parameters comprises the steps of:

3                   encoding the pitch lag from the interpolated pitch lag parameter;  
4                   encoding the pitch gain from the interpolated pitch gain parameter;  
5                   encoding the index of the fixed codebook from analytical forms; and  
6                   encoding the gain from the fixed codebook gain parameter.

1                   31.     The method of claim 28, wherein the step of analysis in the excitation  
2 domain comprises the steps of:

3                   selecting the pitch lag from interpolated pitch lag parameter as an initial value;  
4                   searching the pitch lag in the excitation space;  
5                   searching the pitch gain in the excitation space;  
6                   constructing the target signal for the fixed codebook search;  
7                   searching the fixed codebook index in the excitation space;  
8                   searching the fixed codebook gain in the excitation space; and  
9                   updating the previous excitation vector.

1                   32.     The method of claim 28, wherein the step of analysis in the filtered  
2 excitation space comprises the steps of:

3                   selecting the pitch lag from the interpolated pitch lag parameter as an initial  
4 value;  
5                   searching the pitch lag in the filtered excitation space;  
6                   searching the pitch gain in the filtered excitation space;  
7                   constructing the target signal for fixed codebook search;  
8                   searching the fixed codebook index in the filtered excitation space;

9                    searching the fixed codebook gain in the filtered excitation space; and  
10                    updating the previous excitation vector.

1                    33.     The method of claim 28, wherein the step of analysis in the combined  
2                    excitation space and filtered excitation space comprises the steps of:  
3                            selecting the pitch lag from the interpolated pitch lag parameter as an initial  
4                    value;  
5                            searching the pitch lag in the excitation space;  
6                            searching the pitch gain in the excitation space;  
7                            constructing the target signal for fixed codebook search;  
8                            searching the fixed codebook index in the filtered excitation space;  
9                            searching the fixed codebook gain in the filtered excitation space; and  
10                            updating the previous excitation vector.

1                    34.     The method of claim 28, wherein the excitation mapping strategy is not  
2                    only restricted to the above four strategies, but a combination of these strategies can be  
3                    selected as a new mapping strategy.

1                    35.     The excitation mapping is performed without going back to the speech  
2                    signal domain.

1                    36.     As in claim 1, with the addition of a silence frame transcoding unit  
2                    which can perform rapid conversion of silence frames from one speech coding standard to  
3                    another. This involves mapping the comfort noise parameters.

1                    37.     As in claim 1, but where the quick mapping and tuning engine consists  
2                    of a voice activity detector for generating silence frames. The voice activity detector makes  
3                    its speech/silence decision based on the parameters in the CELP space.

1                    38.     As in claim 1, but with the addition of a mechanism for dynamically  
2                    changing the excitation mapping strategy used, thereby providing a mechanism to adapt to  
3                    available computational resources and allow for graceful quality degradation under load.

1                    39.     A method for processing CELP-based compressed voice bitstreams  
2                    from source codec to destination codec formats, the method comprising:

3                   transferring a control signal from a plurality of control signals from an  
4 application process;  
5                   selecting one CELP mapping strategy from a plurality of different CELP  
6 mapping strategies based upon at least the control signal from the application; and  
7                   performing a mapping process using the selected CELP mapping strategies to  
8 map one or more CELP parameters from a source codec format to one or more CELP  
9 parameters of a destination codec format.

1                   40.     The method of claim 39 wherein the plurality of CELP mapping  
2 strategies includes:

3                   CELP parameters direct space mapping; or  
4                   analysis in the excitation domain; or  
5                   analysis in the filtered excitation domain.

1                   41.     The method of claim 39 wherein the selection of the CELP mapping  
2 strategy for an application is predetermined during a setup process or construction process.

1                   42.     The method of claim 39 further comprising receiving the control signal  
2 at a switching module, the switching module being coupled to each of the plurality of  
3 mapping strategies.

1                   43.     The method of claim 39 wherein the control signal is provided based  
2 upon the computational resource characteristics of the selected CELP mapping strategy.

1                   44.     The method of claim 39 wherein one or more of the plurality of  
2 mapping strategies is provided in a library in memory.

1                   45.     The method of claim 39 further comprising:  
2 encoding one or more CELP parameters for the destination codec; and  
3 processing a destination CELP bitstream by packing one or more CELP  
4 parameters for the destination codec.

1                   46.     The method of claim 39 further comprising transferring the packed  
2 destination CELP bitstream to the destination codec.

1                   47.     A system for processing CELP-based compressed voice bitstreams  
2 from source codec to destination codec formats, the system comprising:

3                   one or more codes for receiving a control signal from a plurality of control  
4 signals from an application process;  
5                   one or more codes for selecting one CELP mapping strategy from a plurality  
6 of different CELP mapping strategies based upon at least the control signal from the  
7 application; and  
8                   one or more codes for performing a mapping process using the selected CELP  
9 mapping strategies to map one or more CELP parameters from a source codec format to one  
10 or more CELP parameters of a destination codec format.

1                   48.     The system of claim 47 wherein the plurality of CELP mapping  
2 strategies includes:

3                   one or more codes directed to CELP parameters direct space mapping; or  
4                   one or more codes directed to analysis in excitation space; or  
5                   one or more codes directed to analysis in filtered excitation space.

1                   49.     The system of claim 47 wherein the selected CELP mapping strategy is  
2 for a predetermined application.

1                   50.     The system of claim 47 further comprising one or more codes directed  
2 to receiving the control signal which is provided to a decision module, the decision module  
3 being coupled to each of a plurality of mapping strategies.

1                   51.     The system of claim 47 wherein the control signal is provided based  
2 upon the computational resource characteristics of the selected CELP mapping strategy.

1                   52.     The system of claim 47 wherein one or more codes directed to the  
2 plurality of mapping strategies are provided in a library in memory.

1                   53.     The system of claim 47 further comprising  
2                   one or more codes directed to encoding one or more CELP parameters for the  
3 destination codec; and

4                   one or more codes directed to processing a destination bitstream by packing  
5 the one or more CELP parameters for the destination codec.

1                   54.     The system of claim 45 further comprising one or more codes directed  
2 to transferring the destination CELP bitstream to the destination codec.

1                    55.     The system of claim 45 further comprising one or more codes directed  
2     to transferring the destination CELP bitstream to a storage location.